1. Introduction

Artificial Intelligence (AI) is the motor that fuels the most profound revolution in the history of humankind. The new era of information society [10], even called “the Fourth Revolution” [8], has produced deep changes into sciences, economies, and societies [9]. This revolution is not only about formal heuristics, or the called “algorithmic society”, but also is explicitly related to the creation of strong interactions between humans and machines. Consequently, these machines must be able to deal with the most intrinsic feature of human nature: emotions. If emotions had been historically considered as disturbing elements of human rationality, the neurological revolution and new experimental tools such as magnetic scanning of the brain (with great success for fMRI) made possible a complete Copernican turn into the evaluation of the role of emotions in cognitive frameworks [13,14]. This process led experts in natural and artificial cognition to consider the necessity of including emotions into their models [37]. Affective computing or social robotics [7,44] increased not only the necessity of studies about how to improve the emotional interaction between humans and machines [57,59] but also how to design cognitive architectures which included biomimetic elements related with emotions [34,58]. The range of emotional aspects with fundamental interest for robot engineers and experts is comprehensive: human–robot interaction, robot task-planning, energy management, social robotics, body design, care robotics, service robotics, among a very long list [2,6,18,20,43]. Practically, there is no field related to robotic and AI which is not directly or indirectly related to the implementation of emotional values. Moreover, embodiment [42] is not a mandatory aspect for considering such emotional elements, because their role embraces fundamental mechanisms of thinking, with a distinctive and relevant role in all those things related with creativity and complex evaluations [12,19,41,53].

Hence, the research investments on emotional robotics have turned from a particular and collateral aspect of robotics and AI studies, to occupy a fundamental and growing area among experts. Emotional or Kansai designs are even becoming standard procedures today [39]. At the same time, there is a huge demand for social robots and intelligent systems, which must also connect with the Internet of Things [3]. The challenges for the understanding, design and implementation of artificial systems which use emotional values for their multimodal and holistic (or general) functioning is, consequently, one of the most critical and fundamental aspects of contemporary researches.
2. Emotional artificial intelligence

In [37] Minsky claims that emotions are just different “ways to think” for addressing different problem types existing in the world, a mean the mind uses to increase our intelligence. He eventually concludes that there is no distinction between emotions and other kinds of thinking. Besides this crucial intuition, the fundamental problem of what is the role of emotions in intelligent behaviour remains open. Whether emotions are evolutionary artefacts or are learned from socialisation and experience, they play the role of particular heuristics, aimed at summarising, focusing and prioritising cognitive tasks. Let think for instance to the role of pain for individual safety, happiness in reinforcement learning, or behavioural responses in non-human animals, e.g. freezing, fleeing, or fighting.

An excellent contribution to modelling emotions came from neurophysiology, which has provided the biological basis for identifying emotional categories. As Barrett states in [4], emotions are socially constructed and have biological evidence.

The seminal works of Ekman on facial emotions [17] has led to the classical model of the six basic categories: anger, disgust, fear, joy, sadness, and surprise, where each real face is assumed to express a weighted combination of such basic pure emotions. The Ekman model, with extensions [16], is also applied to textual emotion recognition. Emotions are physical experiences universal to all humans, and many other animals. Many species can also experience and detect cross-species emotions.

From the computational point of view, a straightforward solution to the task of recognising human-generated emotions is the application of machine learning techniques such as text Semantic Analysis, Naïve Bayesian Networks, Support Vector Machines, Hidden Markov Models and Fuzzy & Neural Networks, to various types of human input with emotional labels [11,32,33,38,50].

The identification of the appropriate features of the input data represents a challenge of increasing complexity, as textual/visual data are massively acquired, and personal devices include sensors which can provide data on physical evidence of emotions.

For instance, when extracting emotions from text, e.g. from social network posts [22,36], the contextual information is prevalent. Indeed, some words can be very much emotionally charged by themselves [5,21,23], but their detection is not sufficient since modifiers and the preceding/following parts of the dialogue significantly influence the resulting relevance and the quality of the detected emotions. For example, an answer given by the single word “yes” can convey different emotions, depending on the particular question asked. Early approaches to speech management often merely consist of generating a text transcript of the speech and applying known text-based techniques. It is apparent that a more accurate context should include both textual and associated acoustic features, such as tone, pauses, voice pitch and expressivity. The inclusion of other non-verbalised features strongly rely on image analysis for detecting facial macro/micro expressions [45], and video analysis for dynamic aspects of facial expressions, gestures recognition, gesture pace/speed, body movements [27].

Another major problem with the supervised machine learning approach is to provide the right labelling. While providing supervised labelling of animals, e.g., dogs, cats, cows appearing in an image is, in most of the cases, a quite straightforward task, the same cannot be done for labelling emotions in faces, voices or behaviours, and assigning them a dimensional quantification [24]. Objective labelling is cumbersome also when dealing with emotions in the text [40]. Tools like WordNet-Affect, an extension to WordNet, where affective concepts labels are attached to affective words, is significantly affected by biases introduced by supervising experts, and by the fact that it does not adequately capture the contribution to emotions given by the context [26]. Moreover, the emotional models are far from being stable; the widely accepted Ekman model itself has been subject to extensions with the introduction of new emotions, and multidimensional models are preferred, but more difficult to compute automatically.

A great opportunity is provided by physical sensors embedded in personal handheld devices, which allow detecting the real physiological evidence of emotion manifestations (e.g., skin temperature/humidity, myoelectrical and electrodermal activity, heartbeat, blood pressure). From the algorithmic point of view, there is an excellent interest in exploring unsupervised labelling approaches based on Neural Networks (NN), in particular, Autoencoders and Convolutional NN (CCN) [48]. The idea is that we somewhat renounce to know, in analytical terms, what the real emotion model is, provided that the model embedded in the trained NN allows us to implement specific operations, such as emotion recognition, computing distance and similarity among emotions, triggering emotion and context driven behaviour.
Symmetric to the issue of recognition is emotion synthesis, i.e. the generation of emotional affordances by an intelligent agent. Emotional affordances are artefacts, i.e. sounds, colours, texts, gestures, behaviours, produced as systems output, which can be percept by an interacting (human) agent as carrying an emotional meaning.

A great effort has been done in this respect, in mimicking human surface behaviours by artificial synthesised characters and avatars, having facial expressions and expressive voice synthesis, as well as making robots reproduce human-like gestures and body motion [57]. There are still a lot of open issues in designing new affordances for non-human-like systems, for instance we can wonder what kind of emotional affordance can provide a vacuum cleaner robot following or respecting his human user, or more simply what kind of non-verbal emotional affordance a web interface can dynamically provide, e.g. by changing colours or shapes.

A crucial issue in managing emotion synthesis is the decision process when to use those emotional affordances in the interactive dialogue and which affordances are more appropriate among alternative ones [25] to the context and the individual user.

A major limitation of current emotion models is that they mostly focus on identifying the basic component emotions, and they tend to locate emotions in short or minimal time intervals.

Future approaches will need to address the issue of abstraction level in emotion models, i.e. to cope with higher level emotional concepts which denote a complex emotional and contextual state, e.g. moods [27], which cannot be reduced merely to a vector of basic emotions, but they express more articulated relationships among emotions, behaviour and context. For instance, an optimistic mood or attitude expresses a temporary state of a subject which tends to privilege certain emotions. Another related aspect worth of investigation is the notion of distribution of emotions over time. In this case, higher-level emotional concepts aim at summarising the emotional content of a series of events, system responses and emotional affordance dialogues with the user. For instance, looking to the concept of user experience, as intended in the Human Machine Interaction area, we notice that it is indeed emotionally characterised. Often, a system-user experience can be described either as difficult, cool, exciting, reliable, easy, boring, seamless, where these labels summarise a unique dynamic distribution of user emotions and systems emotional affordances over the time, and their complex and articulated relationships. The concept of emotional experience mostly relates to time and behaviour, i.e. affordances dialogue, rather than with regular events separated from their history and temporal context.

Another exciting application regards collective emotions, e.g. to model the level of activity of an ethnic group, or estimating stress levels in pedestrian crowds.

3. Emotional robotic machines

As above argued, robots are expected to become capable of perceiving others’ emotions, develop their emotional state and manifest it. Arguably, the scenario where the developments in emotional robots are beginning to occur is home robotics, namely with those robots that are entering our homes in the form of entertainment robots (e.g. robotic animals or puppets) [62] or companion robots (e.g. robots for the elderly) [60] and possibly, also in the form of service providers (e.g. vacuum cleaners) [51]. In all the above contexts, the close and personal interaction with users makes suitable the introduction of emotional components in the robot design and implementation. In particular, the robots that are designed for the children and the elderly, already embody several features of the emotional design. Psychological studies show that for humans, the elderly and children in particular, robots can impersonate missing or imaginary living subjects [47]. In such circumstances, they attribute emotions to them and develop emotions towards them, even independently of the specific emotional design. If this offers another exciting design perspective for emotional robots, on the other hand it raises basic ethical questions on the principles that robot design should respect. As a consequence, the design of emotional robots should be highly intertwined with the ethics of such robots [55].

A specific area of application of emotional design is that of robots that interact with cognitively impaired patients [56]. Under specific therapeutic guidelines, a suitable design of emotional robots can have an impact on the quality of life of the patients, as well as on their rehabilitation [31,46]. Other applications, where the emotional component can have a prominent role are ad-hoc systems for training and education [49] (e.g. dietary robots).

In the home robots, as above described, the manifestation of emotions can be already successfully achieved through multiple modalities, ranging from
appearance to voice and gestures; however, the understanding of human emotions from facial expressions, gesture, spoken language is still somewhat basic, due to the limitations of the hardware, i.e. input sensors.

Other scenarios, where the analysis of human emotions plays a key role, are provided by situations where the human is handling a device that requires full control of emotions (e.g. teleoperating a rescue robot, a drone, driving a car, controlling the motion of a sophisticated industrial device). In particular, this issue is undergoing a fast development in the automotive sector. Systems that monitor the attentive and emotional state of the driver [35] will be installed in cars, even before they will be fully autonomous.

The above-sketched scenarios, where the development of emotional robots is already taking place and likely will be developed in the future, certainly do not cover the whole spectrum of opportunities that arise from the field of emotional robotics. They aim to show that there are already significant impacts and that the emotional sphere is a crucial element in the design of the future generations of robots.

4. Emotional machines: The next revolution

Recent advances in Artificial Intelligence design, Deep Learning techniques, human-friendly Robotics, Cognitive Sciences, claim for a revision of the whole field of Affective Computing and the approaches to the retain of emotional machines. This special issue is devoted to the critical innovations that will pave the way for the next advances in Affective Computing. With the current implementations of AI and robotic systems in new human environments and tasks, the need for a good understanding of the necessary emotional mechanisms involved in such processes is of the utmost importance. The daily interactions between humans and smarter devices are increasing exponentially, and the emotional attachments and relationships with machines are fundamental for a reliable and fruitful interaction.

Such considerations lead to the conclusion that the right approach to the “Next Revolution” must be multidisciplinary.

In this special issue, we are proud to present some particularly exciting contributions to such a view.

From a text-based affective computing point of view, we present a simple sentiment analysis applied in a novel language, introducing the first Dictionary of Kazakh sentiment words [61]. For automated visual face emotion recognition based on micro-expressions, [29] proposes a CNN-based system to processing images streamed in real-time from a mobile device, aiming at helping an impaired user who cannot recognize emotions (e.g., for a visual or cognitive impairment), or a user who has difficulties in expressing emotions (e.g., due to a degenerative pathology), to be assisted. Regarding emotions and attention, [28] presents an audiovisual model for emotion recognition by skin response. A functional data analysis approach for emotion labelling and annotation is proposed in [52], evaluating the variations in annotations across different subjects and emotional stimuli, in order to detect spurious/unexpected patterns, and developing strategies in order to combine these subjective annotations into a ground truth annotation effectively. Another article [1], from the psychology side, can be applied -among others- to the long-standing problem of annotations, tackling the relationship among music, emotions and moral judgement. A psychological approach applied to emotional face recognition is proposed in [15] to track humility from body, face and voice, applied experimentally on politicians’ speeches. Regarding, finally, collective emotions, we propose with [54] a model to measure the level of activity of an ethnic group, and with [30] a realistic model to estimate stress levels in pedestrian crowds.

References


[61] V. Franzoni et al. / Emotional machines: The next revolution

Biographies

Valentina Franzoni, Ph.D. in Engineering for Computer Science at the Department of Computer, Control, and Management Engineering at La Sapienza University of Rome, Italy, is currently a trainer of professors and research fellow at University of Perugia, Italy, where she received her BSc degree and MSc summa cum laude and she has been a senior research assistant (2011–2015), as also at the Department of Computer Science, Hong Kong Baptist University (2012). At Perugia University she worked as a Contract Professor for the Master in Systems and Technologies for Information and Communication Security, and Assistant Professor for Multimedia Systems and Computer Science. Her research interests focus on AI, in which she has a record of invited international speeches, publications and awards since 2008, broadening from emotion recognition to image and context semantics, SNA, usability, and e-learning. She is a member of scientific associations such as AAS, IEEE, AAAC, WIE.

Alfredo Milani is Associate Professor at the Department of Mathematics and Computer Science, University of Perugia, Italy. He has been a visiting associate at the Department of Computer Science, Hong Kong Baptist University for ten years (2008–2018). Alfredo received his degree in Information Science from University of Pisa, Italy in 1987, awarded by Olivetti Inc. company. In 1987, he became a research fellow at the National Research Council Institute for Computational Linguistic, ILC-CNR and consultant of the University of Pisa. In 1990, he obtained the position of tenure researcher at the Department of Mathematics and Computer Science of the University of Perugia, where he has been an Associate Professor since 2001. His research interests involve artificial intelligence and focus on evolutionary algorithms, with applications to planning and web-based adaptive systems. He is the scientific leader of the KitLab research lab at the University of Perugia.
Daniele Nardi is Full Professor (since 2000) at Sapienza University of Rome, Italy, Department of Computer, Control and Management Engineering “A. Ruberti”. He is currently responsible for the Master curriculum in Artificial Intelligence and Robotics. Daniele leads the research laboratory “Cognitive Robot Teams”, addressing different research topics: Cognitive Robotics, Localization, Navigation, Perception, Cooperation in multi-robot systems, Human–Robot Interaction, multimodal interfaces and speech. The scientific and technical achievements are deployed in manifold application domains: Ambient Intelligence and robots to support elderly people, Disaster Response robots to explore and gather information from the environment, Soccer Player robots for RoboCup competitions. Daniele Nardi is an ECAI Fellow and was President of RoboCup Federation (2011–2014). He is currently Principal Investigator of EU projects RoCKEU2, for benchmarking robot competitions and Flourish for the use of robotics for precision agriculture.